



Beyond Effectiveness: Evaluating the Public Health Impact of the WISEWOMAN Program

Rosanne P. Farris, PhD, Julie C. Will, PhD, Olga Khavjou, MA, Eric A. Finkelstein, PhD

Interventions that are effective are often improperly or only partially implemented when put into practice. When intervention programs are evaluated, feasibility of implementation and effectiveness need to be examined. Reach, effectiveness, adoption,

implementation, and maintenance make up the RE-AIM framework used to assess such programs. To examine the usefulness of this metric, we addressed 2 key research questions. Is it feasible to operationalize the RE-AIM framework using women's

health program data? How does the determination of a successful program differ if the criterion is (1) effectiveness alone, (2) reach and effectiveness, or (3) the 5 dimensions of the RE-AIM framework? Findings indicate that it is feasible to operationalize the

RE-AIM concepts and that RE-AIM may provide a richer measure of contextual factors for program success compared with other evaluation approaches. (*Am J Public Health*. 2007;97:641–647. doi:10.2105/AJPH.2005.072264)



EVIDENCE-BASED PUBLIC

health research improves the quality of practice by providing systematic information about tested intervention strategies to public health practitioners.¹ The strongest evidence is often gathered from highly controlled research studies^{1,2} that are designed to test whether a well-defined intervention results in health improvements under ideal conditions. Such studies, referred to as efficacy studies,^{3,4} are designed to eliminate alternative explanations of the causes of the health outcomes of the intervention; consequently, a high degree of experimental control is used. Interventions most worthy of replication in practice are those for which efficacy studies show the strongest association between the intervention and the outcome.⁵

Because they work to improve the health of large populations, public health scientists seek interventions that appeal to the public at large, are effective in practice, and will be adopted rapidly by practitioners. Interventions designed for efficacy studies generally appeal to only the most motivated participants, are less effective when implemented outside of controlled research situations,⁶ and are not easily adopted by practitioners because of their complexity. For a public health scientist, the intervention that warrants replication is the one that has the greatest public health impact, is low-cost, efficient, and feasible to implement in a nonresearch population.

The public health field needs a broad, multidimensional approach

to evaluate interventions. Abrams and colleagues⁷ defined the impact of an intervention as the product of its reach (R) and its efficacy (E), where reach is defined as the percent penetration of the intervention into a defined population. These researchers cited 2 extreme intervention scenarios that could result in zero impact: “(1) a very effective, expensive program (100% efficacy) that fails to attract any clients (0% reach) or (2) a self-help brochure delivered to every smoker (100% reach) that does not work at all (0% efficacy).”^{7(p292)}

Glasgow et al.⁶ expanded the 2-component measure (RE) to a 5-component measure (RE-AIM): reach, efficacy or effectiveness, adoption, implementation, and maintenance. Reach indicates the proportion and representativeness of the target population that participated in the program. Efficacy or effectiveness is the positive program outcomes, minus the negative outcomes. Adoption refers to the proportion and representativeness of settings and people that will adopt the program. Implementation is the extent to which the intervention is implemented as intended. Maintenance is the extent to which the program is sustained over time. The overall public health impact of the intervention is measured by combining all 5 dimensions to create a composite score.

The goal of the Well-Integrated Screening and Evaluation for Women Across the Nation (WISEWOMAN) public health program is to improve the health of midlife, uninsured women by providing cardiovascular screening

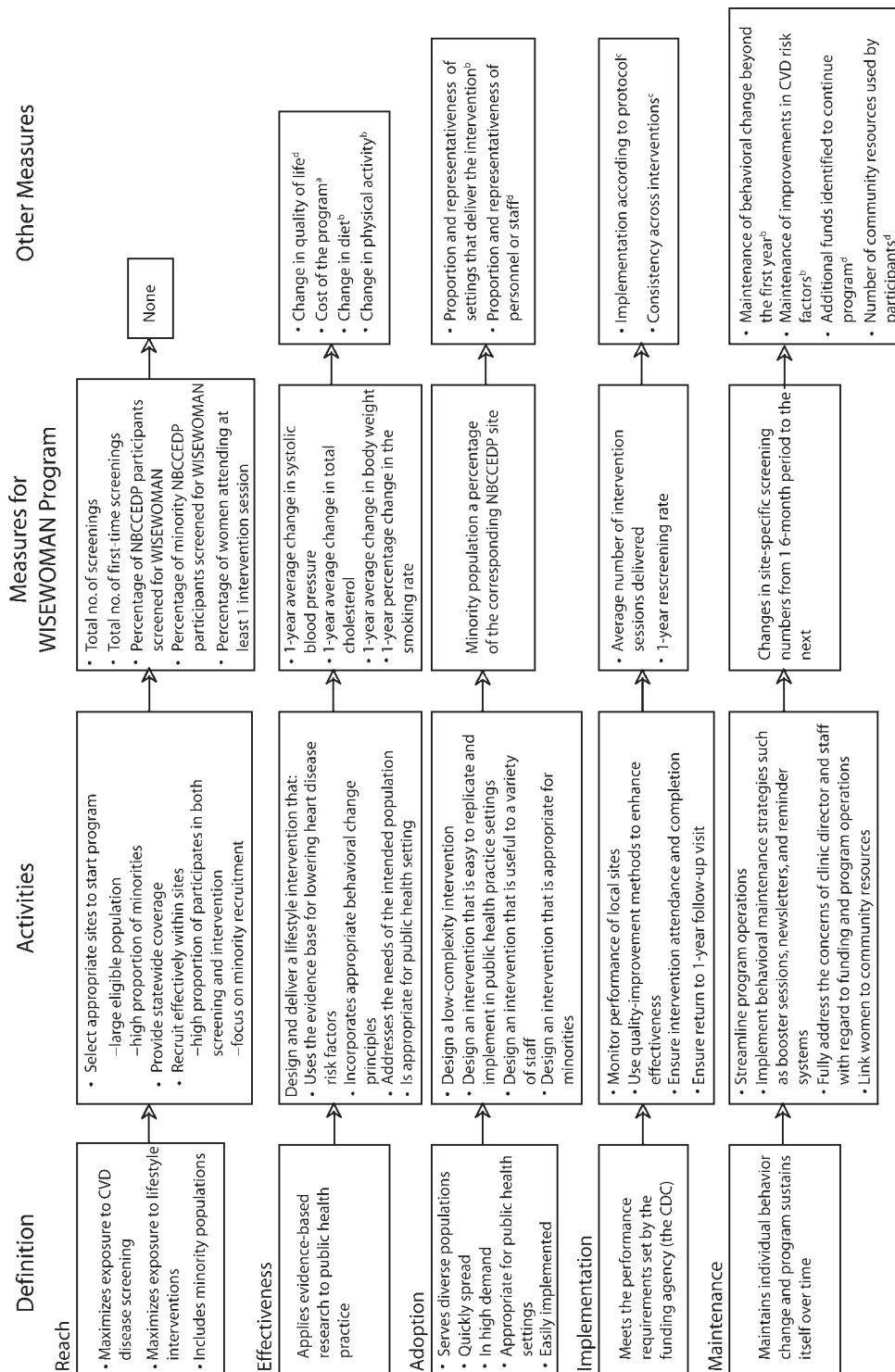
and lifestyle intervention.⁸ As part of the program evaluation effort, data from the 15 projects where the WISEWOMAN program is implemented were used to examine the feasibility and effectiveness of adding a cardiovascular disease prevention component to the National Breast and Cervical Cancer Early Detection Program (NBCCEDP). To assess whether the RE-AIM framework is useful for evaluating WISEWOMAN public health programs, we addressed 2 key research questions: (1) Is it feasible to operationalize the RE-AIM concepts using existing WISEWOMAN program data?¹ and (2) How does the determination of a successful WISEWOMAN program differ if effectiveness alone or a broader approach, such as RE-AIM, is used as a measure?²

METHODS

We used 2001–2003 WISEWOMAN and NBCCEDP data to assess the public health impact of 14 WISEWOMAN sites within North Carolina, which is 1 of the 15 WISEWOMAN projects currently funded by the Centers for Disease Control and Prevention. WISEWOMAN and NBCCEDP collect standard data biannually, including demographic data and physiological measures. Figure 1 illustrates how each of the 5 RE-AIM dimensions for the WISEWOMAN program were operationalized. For each RE-AIM dimension, Figure 1 outlines the components of a successful program, the essential activities, and the public health impact measures available from the WISEWOMAN

data that embody Glasgow et al.’s definitions of each dimension. For example, a successful WISEWOMAN program addresses health disparities by choosing sites that have minority populations. This is measured by the adoption dimension. The implementation dimension measures a successful program on the basis of the performance requirements of the funding agency (i.e., Centers for Disease Control and Prevention). For example, a specified number of intervention sessions should be delivered to each participant. Figure 1 also lists additional public health measures that follow Glasgow et al.’s definitions but are not currently available from existing WISEWOMAN data.

As shown in Figure 1, reach is measured using (1) the total number of screenings, (2) the total number of women screened for the first time, (3) the percentage of NBCCEDP participants screened for WISEWOMAN, (4) the percentage of minority NBCCEDP participants screened for WISEWOMAN, and (5) the percentage of women attending at least 1 intervention session. Effectiveness is measured using 1-year average changes in systolic blood pressure, total cholesterol, body weight, and percentage change in the smoking rate. Racial representativeness of sites that adopt the WISEWOMAN program is measured by assessing the minority population of the corresponding NBCCEDP sites. Implementation assesses whether the program meets requirements of the funding agency and is measured using the average number of intervention sessions delivered



Note. RE-AIM = reach, efficacy and effectiveness, adoption, implementation, and maintenance; WISEWOMAN = Well-Integrated Screening and Evaluation for Women Across the Nation; NBCCEDP = National Breast and Cervical Cancer Early Detection Program.

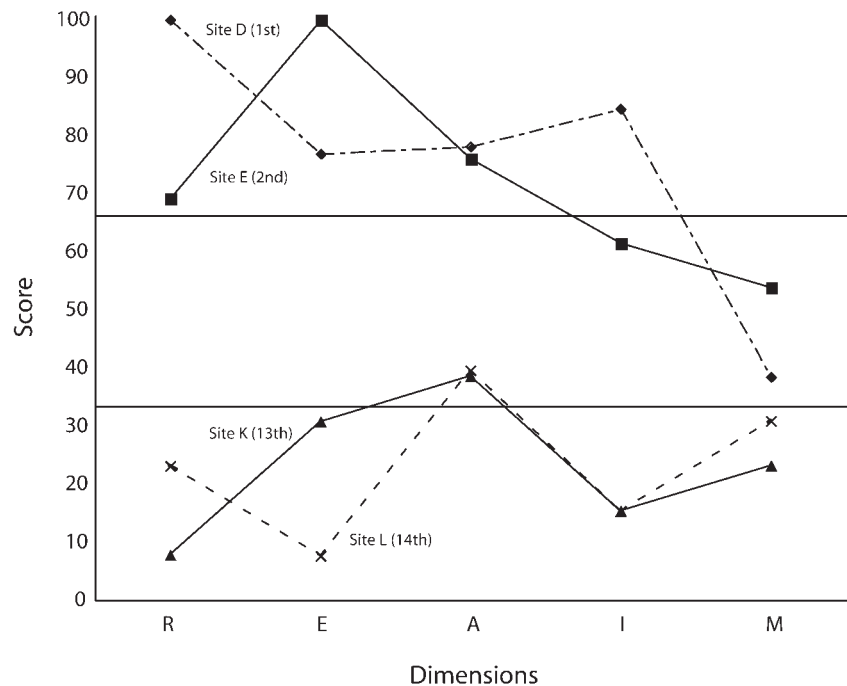
^aData already collected but not available at the site level.

^bPlanned future data collection.

^cAppropriate measure for research but not for public health programs.

^dAppropriate measure for public health programs but ability to collect additional data is limited.

FIGURE 1—Flowchart of the use of the RE-AIM framework for the evaluation of a WISEWOMAN program in general and as a component of NBCCEDP



Note. RE-AIM = reach, efficacy or effectiveness, adoption, implementation, and maintenance; WISEWOMAN = Well-Integrated Screening and Evaluation for Women Across the Nation. Scores fell into 1 of the following tertiles: tertile 1 = 0–33.3; tertile 2 = 33.3–66.6; tertile 3 = 66.6–100.

FIGURE 2—North Carolina WISEWOMAN RE-AIM evaluation framework dimensions: 2 sites with high performance (sites D and E) and 2 sites with low performance (sites K and L).

and 1-year rescreening rates. Maintenance is assessed using changes in site-specific screening numbers from one 6-month period to the next.

To determine which sites were most successful, we calculated the measures outlined in Figure 1 for each of the 14 WISEWOMAN sites in North Carolina. First, the sites were ranked from highest to lowest on each of the 13 measures. Second, the rank scores were averaged for all measures within each dimension. Third, using Glasgow et al.'s suggestion that each of the 5 dimensions be scored on the same scale to allow easy comparison across dimensions, the summary rank score for

each dimension was converted so that it ranged from 0 to 100 (normalized scores). Fourth, each site's RE-AIM dimension scores were plotted on a graph to illustrate differences among sites (Figure 2). Finally, the normalized scores for each dimension were averaged to create a composite RE-AIM score that measures the success of each site. (A detailed, step-by-step outline of the analysis is available as a supplement to the online version of this article.)

We needed to address whether the determination of a successful WISEWOMAN site differs on the basis of using effectiveness alone or the broader RE-AIM approach. To examine this, we compared

each site's ranking on (1) the overall RE-AIM composite score, (2) the effectiveness score alone, and (3) the average of reach and effectiveness (i.e., a modification of Abram's conceptual model).

RESULTS

Table 1 presents the measures constructed for each RE-AIM dimension and the average, lowest, and highest values for each measure. There is considerable performance variation among the 14 sites across RE-AIM dimensions. For the reach dimension, the percentage of the target population (NBCCEDP participants) screened for WISEWOMAN

averaged 39% during 2001–2003; the site that had the lowest performance screened 8% of the target population, and the site that had the highest performance screened 57% of the target population. For the effectiveness dimension, the average reduction in systolic blood pressure after 1 year was 2.7 mm Hg and ranged from a 5.0-mm Hg increase among women at the site with the lowest performance to an 8.0-mm Hg reduction among women at the site with the highest performance. For adoption, the minority population across sites ranged between 3% and 77%. The next step would be to conduct case studies to determine why the variation in sites exists and to identify practices that contribute to high and low performance.

Figure 2 plots the RE-AIM dimension scores for 2 high-performing sites and 2 low-performing sites. This illustrates the relative strengths and weaknesses of each of the sites. RE-AIM dimension rankings contrast considerably between the high- and low-performing sites; sites D and E rank higher for all dimensions than sites K and L. For example, site D has a reach score of 100; whereas, site K has a reach score less than 10. Effectiveness scores are 80 and 100 for the 2 high-performing sites but less than 30 for the low-performing sites. A visual representation of performance, such as Figure 2, will help programs to target areas for improvement at each site.

The results of using the composite RE-AIM scores to rank sites according to their overall public health impact are presented in

**TABLE 1—RE-AIM Measures and Performance Values for 14 WISEWOMAN Sites: North Carolina, 2001–2003**

RE-AIM Dimension	Measure	Site Performance		
		Average	Lowest	Highest
Reach	Total number of screenings ^a	236	100	480
	Total number of first-time screenings	143	60	346
	NBCCEDP participants screened for WISEWOMAN, ^b %	39	8	57
	Minority NBCCEDP participants screened for WISEWOMAN, ^c %	41	6	67
	Women attending at least 1 intervention session, %	94	80	100
Efficacy or effectiveness	1-year average change in systolic blood pressure, mm Hg	-2.7	5.0	-8.0
	1-year average change in total cholesterol, mg/dL	1.5	12.0	-11.0
	1-year average change in body weight, lb	-0.29	9.0	-4.0
	1-year change in the smoking rate, %	-17	0	-50
Adoption	Minority population of the corresponding NBCCEDP site, %	48	3	77
	Average number of intervention sessions attended	1.4	1.0	3.0
Implementation	1-year rescreening rate, ^d %	29	9	64
Maintenance	Changes in site-specific screening numbers from one 6-month period to the next (percentage change from the first period to the last), %	76	-43	586

Note. RE-AIM = reach, efficacy or effectiveness, adoption, implementation, and maintenance; NBCCEDP = National Breast and Cervical Cancer Early Detection Program; WISEWOMAN = Well-Integrated Screening and Evaluation for Women Across the Nation.

^aTotal number of screenings is based on budget allocated to sites and may not be the best measure for North Carolina; however, this is not the case for most projects.

^bOnly women enrolled in NBCCEDP are qualified for the WISEWOMAN program; therefore, NBCCEDP participants are the target population.

^cMinority is defined as non-White women of known race/ethnicity.

^dRescreening rates are influenced by changes in health insurance status and income changes that could affect program eligibility.

Table 2. A rank of 1 indicates the site that has the highest performance, and a rank of 14 indicates the site that has the lowest performance. The top tertile of sites (those ranked from 1 through 5) across the 3 methods (i.e., effectiveness alone, reach and effectiveness, RE-AIM) have been noted. RE-AIM framework rankings are compared with the other rankings using effectiveness alone and the average of the reach and effectiveness dimensions.

Table 2 illustrates how the use of the different evaluation methods can affect the selection of the top 5 high-performing sites for case studies in evaluation research. If effectiveness alone is the sole determinant of high performance, sites C, D, E, H, and J would

be selected. If RE-AIM is used as the determinant of performance, sites A, D, E, G, and N would be selected. Sites D and E would be among the 5 top-performing sites regardless of which method is used to rank the sites. However, sites G and N would be selected only if the RE-AIM framework is used as the determinant. Site J would be selected if effectiveness alone is used to rank sites. Thus, sites would be selected differently for case studies of program practices if RE-AIM were used as the determinant of performance instead of effectiveness alone.

DISCUSSION

Using existing program data from WISEWOMAN and the

NBCCEDP, we successfully operationalized the 5 dimensions of the RE-AIM model and identified high- and low-performing sites. This task was not without challenges. We needed to decide how data that are routinely collected could be applied to the 5 RE-AIM dimensions. We had multiple measures for some dimensions, but measures for other dimensions were more difficult to identify. We also needed to convert measures to similar units so they could be combined to generate an overall score. This required the use of ranks, which was not part of Glasgow et al.'s methodology but was accepted by our panel of experts. To better understand the relationships among the 5 RE-AIM dimensions, we calculated

correlations between each of the measures (available from the authors). We found positive correlations between effectiveness and implementation (0.45), reach and adoption (0.33), and adoption and implementation (0.24), and a negative correlation between implementation and maintenance (-0.25). None of the correlations was statistically significant. These results suggest that each measure provides additional information concerning the overall benefits of the program.

The broader dimensions included in the RE-AIM framework (rather than just the effectiveness) are important contributors to public health impact. Several evaluation experts have addressed this issue. Nutbeam,⁹ for example, described intervention program evaluation as a “complex enterprise” and suggested that changes in outcomes should not be the only standard for a successful program. This concept of “effectiveness” includes measures such as changes in knowledge and skills at the individual level, social action and changes in social norms, and changes in policy and organizational practices as a result of the intervention. Nutbeam suggested that decisionmaking for evidence-based practice in health promotion should be based on the best available evidence concerning intervention program effectiveness and the intervention program's application in real-life circumstances.¹⁰

In addition, Green¹¹ raised the following question: “Where did the field get the idea that evidence of an intervention's efficacy from carefully controlled trials could be



TABLE 2—Results Using the Composite RE-AIM Scores to Rank WISEWOMAN Sites According to Their Overall Affect on Public Health Across 3 Different Methods: North Carolina, 2001–2003

Site	Method		
	RE-AIM	Effectiveness	Reach and Effectiveness
A	3 ^a	6	3 ^a
B	10	9	12
C	8	2 ^a	5 ^a
D	1 ^a	4 ^a	1 ^a
E	2 ^a	1 ^a	2 ^a
F	11	8	9
G	4 ^a	11	8
H	6	5	4 ^a
I	7	14	11
J	9	3 ^a	6
K	14	10	13
L	13	13	14
M	12	12	10
N	5 ^a	7	7

Note. RE-AIM = reach, efficacy or effectiveness, adoption, implementation, and maintenance; WISEWOMAN = Well-Integrated Screening and Evaluation for Women Across the Nation.

^aSites ranked in the top tertile (1–5).

generalized as the ‘best practice’ for widely varied populations and situations?”^{11(p167)} Green discussed the need to consider interventions in the context of the social and cultural, economic, and occupational circumstances of the individual as well as the target group and organizational variations in their many combinations within populations. He suggested that preoccupation with internal validity—the degree to which the observed changes can be attributed to the effect of the intervention—in evidence of effectiveness from research studies causes external validity—the degree to which the findings can be generalized to other settings or populations—to receive little attention in final recommendations of

best practices. The Centers for Disease Control and Prevention *Guide to Community Preventive Services*² emphasizes that the strength of evidence for the effectiveness of population-based interventions should be linked to recommendations for population-based and public health interventions. However, the Centers for Disease Control and Prevention has a set of procedures for considering applicability to local situations. First, it assesses in which populations and settings the interventions were studied, and then it determines whether these populations and settings are representative of other populations and settings of interest. Green called for a more systematic study of place, organizational settings, social

circumstances, and culture as part of the research agenda to guide health promotion practice.

Other researchers^{12,13} have called for a broader definition of impact because of the relative dearth of published studies in the health promotion field and because the effectiveness of health promotion programs relies heavily on how well the program fits within the local contexts. Abrams et al.⁷ defined the impact of an intervention as the percentage of the population that receives the intervention multiplied by the intervention’s efficacy. Glasgow et al.⁶ conceptualized the public health impact of an intervention as a function of 5 factors that are compatible with a social–ecological theory, systems-based approach, and community-based and public health interventions. In so doing, he included external validity factors that affect program success and expanded Abrams et al.’s concept of reach by including the representativeness of the population that receives the intervention. Glasgow et al.’s definition of efficacy (or effectiveness) goes beyond biological outcomes, such as disease risk factors, and includes behavioral outcomes for intervention staff and participants. In addition, he included negative outcomes measured through changes in the participant’s quality of life. The organizational-level components include the proportion and representativeness of settings that adopt the intervention as well as barriers to adoption, the extent to which the program is delivered as intended (implementation), and program-level measures of institutionalization (maintenance).

Glasgow et al.’s evaluation method addresses the conceptual issues of the interventions being studied and recognizes the complexity of determinants of program success, which gives decisionmakers more complete information on which to base program decisions.

There are several limitations to this analysis. First, an intervention program evaluation that uses broad frameworks should be designed before the intervention program is initiated. Because the WISEWOMAN program was initiated in 1995, the RE-AIM framework, which was first presented in 1999, could not be incorporated. WISEWOMAN was designed to include in its evaluation both the reach and effectiveness dimensions from the outset. The social–ecological model was the theoretical construct for the intervention, but because there were no measures to capture adoption, implementation, and maintenance from the program onset, the existing data had to be retrofit to capture these dimensions. The resulting measurements of these 3 dimensions are not optimal. For example, we had less-than-ideal measurements of program fidelity and so could not verify that the program was implemented as intended in every site. However, we plan to develop a method for collecting and reporting this information as part of an ongoing best practices study. The information from this ongoing study, coupled with the RE-AIM measures, can assess whether certain adaptations are appropriate (e.g., no change or improved effectiveness) or inappropriate (e.g., loss of effect). A second limitation was that the



broader frameworks were tested (in 14 sites) in only 1 of the 15 WISEWOMAN projects. Whether the lessons learned in this single project can be generalized to all of the WISEWOMAN projects has yet to be determined. Third, cost and cost-effectiveness are important factors in program evaluation, but it is not clear how these metrics should be incorporated into the RE-AIM framework. Glasgow et al. suggested that cost-effectiveness and cost-benefit are appropriate outcomes and that a population-based, cost-effectiveness index could be calculated by dividing the public health impact (RE-AIM score) by the total societal costs of a program. We are conducting cost-effectiveness studies of each of the 15 WISEWOMAN projects and could incorporate this into the model in the future. Fourth, even though RE-AIM provides a comprehensive framework for program evaluation, other factors influence the success of the program, such as effect of the natural, social, or constructed environments on key program outcomes. Other evaluation frameworks, such as Health Impact Assessment¹⁴ or the Precede-Proceed Model,¹⁵ may include these factors; however, we chose to use RE-AIM because it is more appropriate for evaluation of a behavioral change intervention.

Finally, we are currently unable to assess the validity and robustness of the chosen RE-AIM measures. We will, however, explore predictive validity of our measures to identify high- and low-performing sites by conducting interviews with key program informants (e.g., local coordinators,

project directors, and managers) and assessing whether our results are consistent with their perceptions and experiences. A preliminary analysis revealed that the project manager's subjective response to which 4 sites were high performing and low performing matched the results generated from RE-AIM.

In conclusion, we used WISEWOMAN program data to examine the feasibility of measuring each of the 5 dimensions of RE-AIM and compared this evaluation method to other methods for determining program success. The findings indicate that RE-AIM captures important organizational dimensions not captured by other metrics. These dimensions may be particularly useful for public health practitioners who conduct program evaluation research or monitor program performance. Using the RE-AIM framework when planning and designing intervention programs may lead to the development of studies and programs that have greater public health impact and will enhance translation of evidence-based interventions and their dissemination in the real world setting.¹⁶ Additional research is needed to better define and maximize the impact on public health. However, this investigation has demonstrated that broader evaluation frameworks, such as reach and effectiveness and RE-AIM, contribute more than effectiveness alone. ■

About the Authors

Rosanne P. Farris and Julie C. Will are with the Centers for Disease Control and Prevention, Division for Heart Disease and Stroke

Prevention, Atlanta, Ga. Olga Khavjou and Eric A. Finkelstein are with RTI International, Health, Social, and Economics Research, Research Triangle Park, NC.

Requests for reprints should be sent to Rosanne P. Farris, PhD, Division for Heart Disease and Stroke Prevention, Centers for Disease Control and Prevention, 4770 Buford Highway, MS K-47 Atlanta, GA 30341 (e-mail: rif6@cdc.gov).

This article was accepted December 7, 2005.

Contributors

R.P. Farris and J.C. Will conceptualized the study and helped to interpret the findings, and write and review drafts of the article. O. Khavjou assisted with the study, conducted the data analysis, and reviewed drafts of the article. E.A. Finkelstein assisted with writing the article, interpreting findings, and reviewing drafts of the article.

Acknowledgments

This work was funded by the Centers for Disease Control and Prevention (grant 200–97–0621).

The authors thank Russell Glasgow for his consultation and helpful comments, Carolyn Townsend, Project Director, and other North Carolina WISEWOMAN staff for their valuable suggestions.

Note. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the funding agency.

Human Participant Protection

Collection of WISEWOMAN minimum data elements used in this analysis was approved by RTI's institutional review board.

References

- Truman BI, Smith-Atkin CK, Hinman AR, et al. Developing the Guide to Community Preventive Services—overview and rationale. *Am J Prev Med*. 2000;18(1S):18–26.
- Briss PA, Zaza S, Pappaioanou M, et al. Developing an evidence-based Guide to Community Preventive Services—methods. *Am J Prev Med*. 2000;18(1S):35–43.
- Cochrane AL. *Effectiveness and Efficacy: Random Reflections on Health Services*. London, England: Nuffield Provincial Hospitals Trust; 1971.
- Flay BR. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Prev Med*. 1986;15:451–474.
- Greenwald P, Cullen J. The scientific approach to cancer control. *CA Cancer J Clin*. 1984;34:328–332.
- Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. 1999;89:1322–1327.
- Abrams DB, Orleans CT, Niaura RS, Goldstein MG, Prochaska JO, Velicer W. Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: a combined stepped care and matching model. *Ann Behav Med*. 1996;18:290–304.
- Will JC, Farris RP, Sanders CG, Stockmyer CK, Finkelstein EA. Health promotion interventions for disadvantaged women: overview of the WISEWOMAN projects. *J Womens Health*. 2004;13:484–502.
- Nutbeam D. Achieving “best practice” in health promotion: improving the fit between research and practice. *Health Educ Res*. 1996;11:317–326.
- Nutbeam D. The challenge to provide “evidence” in health promotion. *Health Promot Int*. 1999;14:99–101.
- Green LW. From research to “best practices” in other settings and populations. *Am J Health Behav*. 2001;25:165–178.
- Wandersman A. Community science: bridging the gap between science and practice with community-centered models. *Am J Community Psychol*. 2003;14:227–242.
- Rada J, Ratiama M, Howden-Chapman P. Evidence-based purchasing of health promotion: methodology for reviewing the evidence. *Health Promot Int*. 1999;14:177–187.
- Scott-Samuel A, Birley M, Ardern K. *The Merseyside Guidelines for Health Impact Assessment*. 2nd ed. Liverpool, England: The International Health IMPACT Assessment Consortium; 2001.
- Green L, Kreuter M. *Health Program Planning: An Educational and Ecological Approach*. 4th ed. New York, NY: McGraw-Hill; 2005.
- Klesges L, Estabrooks P, Dzawalowski D, Bull S, Glasgow R. Beginning with the application in mind: designing and planning health behavior change interventions to enhance dissemination. *Ann Behav Med*. 2005;25(special suppl):66–75.